



Chief Marketing Officer

Main analysis

- Customer clustering and segmentation
- Identify the customer segment that drives the most revenue

KPI	Description	Availability	Assumptions
Recency	Days since last customer transaction	Yes	Today = 2015/01/01
Frequency	Number of transactions in the last 24 months	Yes	X
Monetary Value	Total spend in the last 24 months	Yes	X
Customer Lifetime Value	Total amount of money a customer is expected to spend in your business during their lifetime. Granular CLV = (Avg. Revenue per purchase * Avg. Frequency * Profit Margin) * Avg. Lifespan	Yes	Avg. Customer lifespan = 3 y

Customer Data & Features

- KMeans clustering
- KPI performance for each segment
- Marketing strategy for each customer segment

RFM calculated from transaction data

CLV calculated from transaction data

Product family/dep. the customer spent the most money on

Customer ID	Recency	Frequency	MonetaryValue	CLV	Gender	Marital Status	Homeowner	Children	Income	Product Family	Product Department	State or Province	Country
3	275.0	2.0	58.60	2430.70	F	M	True	1	60000.0	Non-Consumable	Health and Hygiene	WA	USA
8	91.0	8.0	129.44	1342.28	M	M	True	2	60000.0	Food	Produce	DF	Mexico
9	52.0	4.0	103.32	2142.83	M	M	True	5	20000.0	Drink	Beverages	BC	Canada
20	132.0	6.0	118.54	1638.99	F	S	True	2	20000.0	Food	Frozen Foods	CA	USA
24	49.0	4.0	17.40	360.87	F	S	False	0	40000.0	Food	Snacks	DF	Mexico
...
10271	140.0	6.0	55.58	768.48	F	S	False	0	80000.0	Food	Produce	Yucatan	Mexico
10273	224.0	8.0	103.58	1074.11	F	S	False	3	80000.0	Food	Baking Goods	WA	USA
10274	331.0	2.0	23.04	955.69	M	S	True	4	60000.0	Non-Consumable	Health and Hygiene	CA	USA
10277	43.0	4.0	44.48	922.50	M	M	False	4	100000.0	Non-Consumable	Household	OR	USA
10278	254.0	2.0	11.96	496.09	F	M	False	0	40000.0	Non-Consumable	Household	BC	Canada

5290 rows x 13 columns

Customer Cluster Analysis: KMeans (k=4)

High Income Homeowner Buyer

- high income / homeowner
- visit less frequent (due to work)



Observations  820 (15.50%)

- **Income** is in average **107% greater** : mean of 120k against 55841 globally
- **86.10%** of the cluster has **true** for **Homeowner** (against 59.87 % globally)
- **Frequency** is in average **20.54% smaller** : mean of 4.13 against 5.20 globally

Young Buyer

- low income, few or no children, low spending in store
- 1/3 people lives in California



Observations  1939 (36.65%)

- **Children** is in average **53.35% smaller** : mean of 1.17 against 2.52 globally
- **Income** is in average **22.14% smaller** : mean of 43476 against 55841 globally
- **MonetaryValue** is in average **30.19% smaller** : mean of 47.28 against 67.72 globally

Big Family Buyer

- Customer who has big family (avg. 3.84 children)
- not living in urban area




Observations  1951 (36.88%)

- **Children** is in average **52.71% greater** : mean of 3.84 against 2.52 globally
- **Income** is in average **21.67% smaller** : mean of 43742 against 55841 globally
- **Frequency** is in average **25.98% smaller** : mean of 3.85 against 5.20 globally

Frequent Buyer

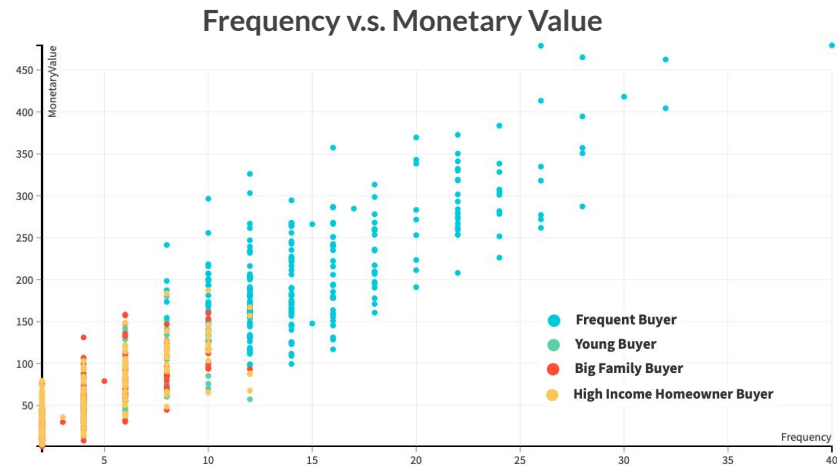
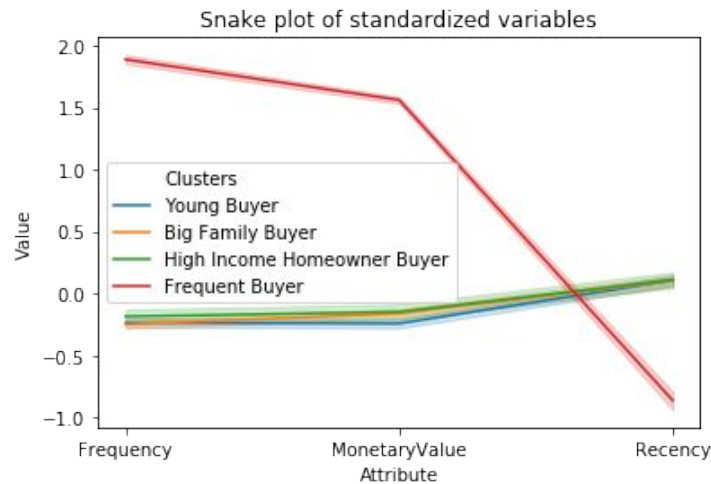
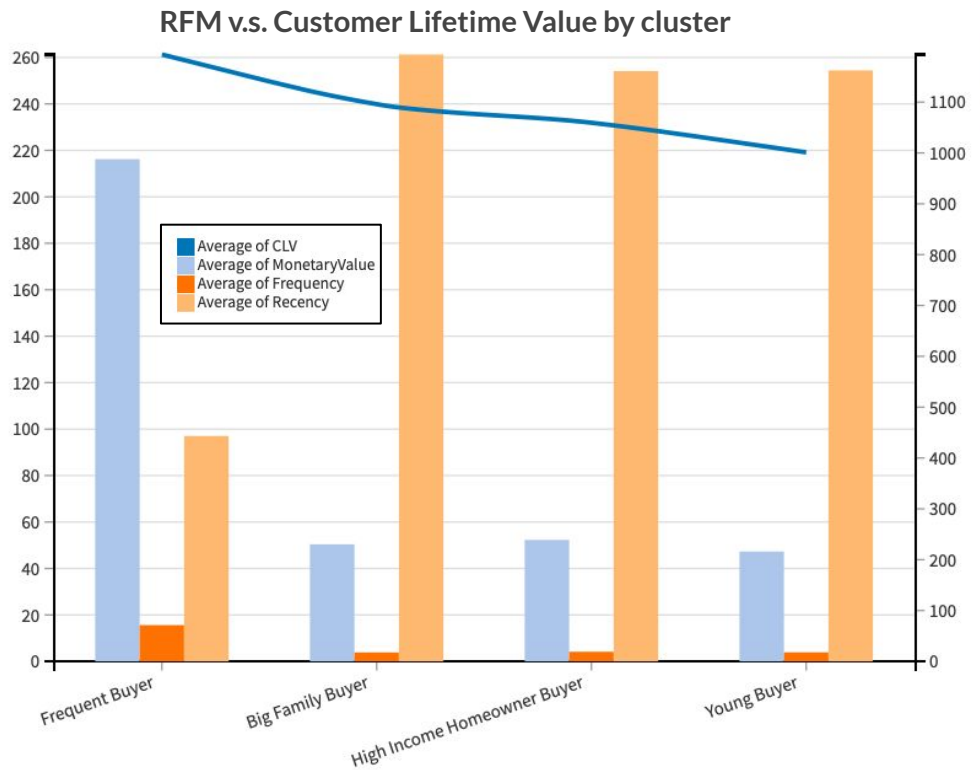
- visit frequent, 91% of the cluster spend the most money on food
- 45% of the cluster lives in WA state



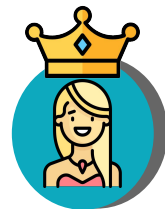
Observations  580 (10.96%)

- **Frequency** is in average **200% greater** : mean of 15.61 against 5.20 globally
- **MonetaryValue** is in average **219% greater** : mean of 216 against 67.72 globally
- **44.83%** of the cluster has **Mexico** for **Country** (against 17.41 % globally)

Cluster KPI Performances

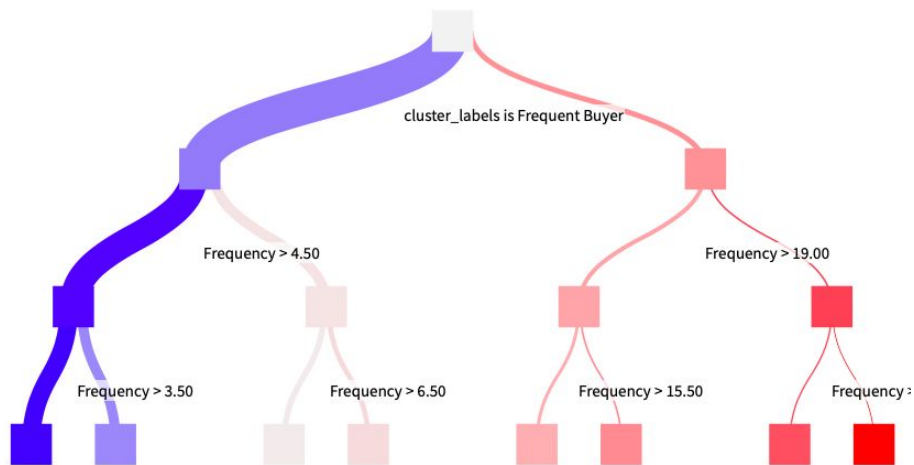


Which customer segment drives the most revenue?

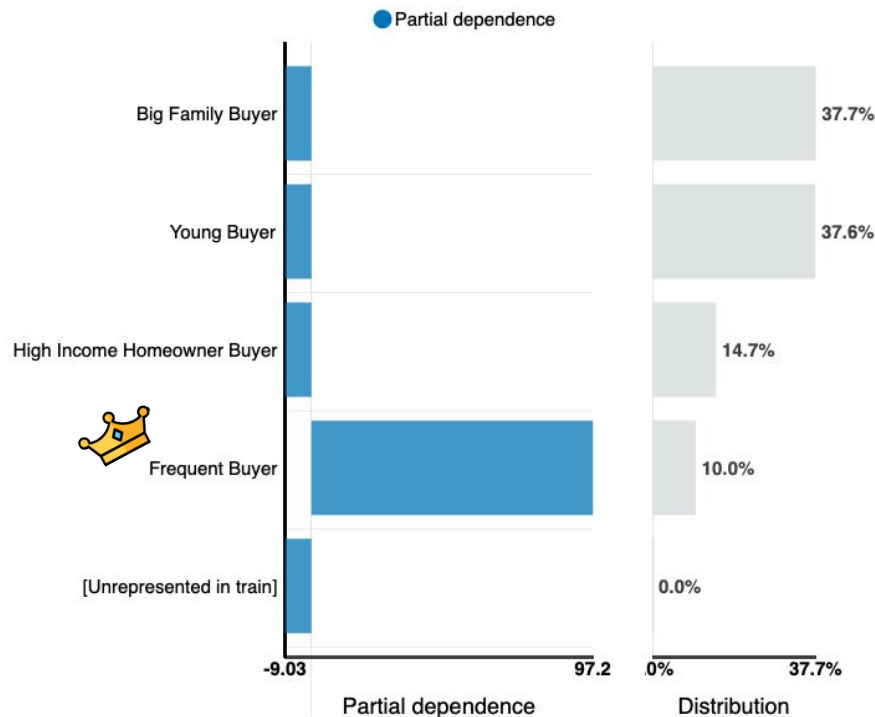


Frequent Buyer

- Supervised Learning - Regression Problem
(Target: Monetary Value)
- Algorithm: Gradient Boost Trees (R2 score: 0.95)
- Conclusion: **Whether the customer is in Frequent Buyer cluster** is the most important variable



Gradient Boost Trees (max tree depth = 3)





CMO: Marketing Strategy

- Marketing campaigns focusing on **Frequent Buyer** cluster (11% of the customer)
 - focus on the stores in **Washington State and Mexico (especially Zacatecas)**
 - discounts/vouchers/marketing campaign on **Food products**
- Marketing Strategy for **other customer segments**:
 - **Family Buyer**: focus on the stores in **rural area**, discounts/vouchers/marketing campaign on **products for children/family**
 - **High Income Buyer**: focus on the stores in **cities (ex: SF)**, marketing campaigns to **stimulate higher frequency**
 - **Young Buyer**: focus on the stores in **California**, discount/campaign on **drinks**